# MINOR STUDIES FROM THE PSYCHOLOGICAL LABORATORY OF YALE UNIVERSITY

Communicated by JOHN E. ANDERSON

## I. AN EXPERIMENT IN TIME ESTIMATION USING DIFFERENT INTERPOLATIONS

By LLEWELLYN T. SPENCER

The problem of the perception of time has been given frequent and almost continuous treatment. Much has been done, especially in Germany, in regard to difference limina, increased uncertainty with increased length of interval, etc. In this country, among other investigators, R. McDougall[1] with thirty subjects found sex differences. He found that women are less accurate than men in their estimates of time-intervals and had a greater tendency to overestimate. Yerkes and Urban[2] conducted a group experiment on a far larger scale, in order to study this difference. Four intervals were used, 18, 36, 72, and 108 sec. The subject recorded his estimation of the interval in seconds. The interval was given by the signals "start" and "stop" spoken by the experimenter. With each of the intervals four different "fillings" or interpolations were used. These were idleness, reading (by the experimenter), writing (at dictation), and estimation (any method except the use of a timepiece). In the experiment approximately two hundred and fifty males and approximately two hundred and fifty females were used. This made a grand total of over 8,400 observations. Of all the observations for all the intervals and interpolations 2.39% were correct for males and 1.05% for females. This bore out McDougall's findings. They also found that the number of correct judgments decreases with the increase in the length of the interval. But upon examining their results, they discovered that in the judgments reported the final digit was most often a multiple of five; 65.9% of the male judgments and 80% of the female judgments ended in 5 or 0. This frequency is too great to be due merely to chance. They also found that even numbers were more frequent than odd numbers as the final digit. The order of decreasing frequency for the last digit was 0, 5, 8, 2, etc. In other words judgments "were strongly influenced by the thought of the conventional time unit, the minute, for in all quantitative work there are errors in favor of the standard of measurement and simple fractions thereof. . . . The longer an interval . . . the more frequently it is judged as the same as the unit or a simple fraction of that unit."[3] Inasmuch as the intervals used were not multiples of 5, it is clear that a certain influence not present in normal behavior affected the judgments of the subjects in the investigation of Yerkes and Urban. The present investigation was undertaken to determine whether a method of judgment which is not

[1] *Science*, N. S. 19, 1904, 708-709.
[2] Harvard Psychological Studies, 2, 1906, 405-430.
[3] *Op. cit.*, 417 and 418.

subject to this influence would produce results varying significantly from those obtained by Yerkes and Urban. The method chosen was that of reproduction.

The interval was given by a sharp rap on the table with a pencil or ruler as a signal of the beginning and of the end of the interval. The subject immediately reproduced the interval by starting and stopping a stop-watch, which was enclosed in a felt-lined case to prevent any reference to the rhythmical ticking. In this way the judgment was freed of any prejudice from "the thought of the conventional time unit." The judgments could be obtained to the fifth of a second. Four intervals were used, 15, 30, 60, and 100 sec. These closely approximate those used by Yerkes and Urban. The experiment was divided into two parts. In the first part the four intervals were given in mixed order, without any interpolation. This corresponds to the "filling" called "estimation" by Yerkes and Urban and will hereafter be designated by that name. The subjects were instructed to reproduce the interval immediately after the final rap, without using any

## TABLES OF RESULTS

TABLE I.
15 SEC.

Estimation

| Sec. | Cases |
|---|---|
| 10–14 | 12 |
| 15–19 | 43 |
| 20–24 | 17 |
| 25–29 | 11 |
| 30–34 | 5 |
| 35–39 | 4 |
| 40–44 | 3 |
| 45–49 | 1 |
| 50–54 | 1 |
| 55–59 | 2 |
| 60–64 | 1 |

TABLE II.
30 SEC.

Estimation

| Sec. | Cases |
|---|---|
| 20–24 | 9 |
| 25–29 | 18 |
| 30–34 | 22 |
| 35–39 | 17 |
| 40–44 | 8 |
| 45–49 | 11 |
| 50–54 | 5 |
| 55–59 | 4 |
| 60–64 | 2 |
| 65–69 | 1 |
| 70–74 | 1 |
| 75–79 | 1 |
| ..... | .. |
| 105–109 | 1 |

TABLE III.
60 SEC.

Estimation

| Sec. | Cases |
|---|---|
| 30–34 | 1 |
| 35–39 | 2 |
| 40–44 | 2 |
| 45–49 | 7 |
| 50–54 | 9 |
| 55–59 | 27 |
| 60–64 | 11 |
| 65–69 | 7 |
| 70–74 | 7 |
| 75–79 | 6 |
| 80–84 | 7 |
| 85–89 | 1 |
| 90–94 | 1 |
| 95–99 | 4 |
| 100–104 | 2 |
| 105–109 | 1 |
| 110–114 | 1 |
| 115–119 | 1 |
| 120–124 | 1 |
| ..... | .. |
| 150–154 | 1 |
| 165–169 | 1 |

TABLE IV.
100 SEC.

Estimation

| Sec. | Cases |
|---|---|
| 35–39 | 1 |
| 40–44 | .. |
| 45–49 | 1 |
| 50–54 | 2 |
| 55–59 | 1 |
| 60–64 | 3 |
| 65–69 | 1 |
| 70–74 | 4 |
| 75–79 | 2 |
| 80–84 | 9 |
| 85–89 | 7 |
| 90–94 | 7 |
| 95–99 | 11 |
| 100–104 | 7 |
| 105–109 | 8 |
| 110–114 | 8 |
| 115–119 | 7 |
| 120–124 | 2 |
| 125–129 | 4 |
| 130–134 | 1 |
| 135–139 | 2 |
| 140–144 | 2 |
| 145–149 | 3 |
| 150–154 | 1 |
| 155–159 | 2 |
| 160–164 | 1 |
| 170–174 | 2 |
| 200–204 | 1 |

method of rhythmic comparison; that is, counting in the stimulus and response intervals, humming tunes, or thinking poetry, and thus gaining a basis for comparison, were forbidden. During the reproduction the subject was idle both in the first set of observations and in the second set, described below. Ten subjects were used throughout the experiment. These were men of the undergraduate body of Yale College and the Sheffield Scientific School and of the Yale Psychological Department. The four intervals were given twice at each sitting and only one sitting was held each day. No attempt was made to keep the sittings regular, but each subject had five sittings on each interval or interpolation, which makes a total of one hundred judgments on each interval or interpolation. As the subjects were ignorant of the success of their judgments, practice-effects could have small influence. Tables I, II, III, and IV give the distribution of judgments for these four intervals in classes of five.

In the second part of the experiment, the 30 sec. interval was used with the following interpolations: Reading of prose by the experimenter, Reading of prose by the subject, Reading of poetry by the experimenter, Reading of poetry by the subject, and Dictation. The interpolation was of course used only in the stimulus interval, and during the reproduction interval the subject spent the time in idleness as in the period of estimation described above. For the prose, selections from current periodicals were used and familiar and especially rhythmical poems were employed in the intervals devoted to poetry. Dictation was from the material of the reading. Table V gives the distribution of judgments for these five interpolations. For convenience of reference, the 30 sec. interval with estimation (idleness) is included also in this table.

TABLE V

INTERPOLATIONS FOR 30 SECS.

| Judgment | Estimation | Prose By Experimntr | Prose by Subject | Poetry By Experimntr | Poetry by Subject | Dictation |
|---|---|---|---|---|---|---|
| 5–9 | | | | | | 2 |
| 10–14 | | | 1 | 1 | | 10 |
| 15–19 | | 5 | 1 | | 2 | 9 |
| 20–24 | 9 | 4 | 6 | 9 | 8 | 10 |
| 25–29 | 18 | 18 | 15 | 5 | 10 | 14 |
| 30–34 | 22 | 24 | 16 | 22 | 16 | 18 |
| 35–39 | 17 | 11 | 19 | 18 | 15 | 8 |
| 40–44 | 8 | 14 | 12 | 24 | 14 | 9 |
| 45–49 | 11 | 9 | 7 | 7 | 13 | 11 |
| 50–54 | 5 | 7 | 11 | 6 | 7 | 3 |
| 55–59 | 4 | 6 | 6 | 3 | 7 | 2 |
| 60–64 | 2 | | 2 | | 2 | |
| 65–69 | 1 | 1 | 1 | 3 | 1 | 1 |
| 70–74 | 1 | | 2 | 1 | 2 | 2 |
| 75–79 | 1 | | | | | |
| 80–84 | | | 1 | | 1 | |
| 85–89 | | | | | 1 | 1 |
| 90–94 | | | | | 1 | |
| 95–99 | | 1 | | | | |
| 100–104 | | | | 1 | | |
| 105–109 | 1 | | | | | |

Of the nine hundred observations, thirty-five are correct. Responses were read to the nearest second, under the assumption that responses within two-and-one-half fifths either side of the unit second were equally distributed within those limits. This gives a percentage of 3.88% correct. Yerkes and Urban found the percentage to be 2.39%, as stated above.

The effect upon the percentage of correct responses of the prejudice in favor of the multiples of 5 can be approximately calculated. It is evident that this prejudice increases the returns for two of the possible final digits, at the expense of the other eight. It is assumed that this prejudice operates in a fairly regular manner. Such an assumption is always made in the description of any psychological trait. For example, we assume that every time the Müller-Lyer illusion is presented there will be a constant tendency to misjudge the distances. The percentage given by Yerkes and Urban of 65.9% of the responses ending in 5 or 0 represents the expected returns for those digits *increased* by a certain percentage at the expense of the other digits. Of all the responses only 20% should end in 5 or 0, on the basis of pure chance. Hence the difference between the 65.9% obtained and the 20% expected, viz., 45.9%, represents an increase of the returns for 5 or 0, due to the disturbing influence. This percentage should be redistributed among the other eight digits, so that $45.9\% \div 8 = 5.73\%$ approximately should be allotted to the percentage of cases of the other final digits. I say 'approximately,' for it was shown in their article that the other final digits were also affected by the prejudice, but not by any means to such a great degree. But this addendum of 5.73% refers only to the final digit in each case. We are interested in the correct tens digit as well. Only a portion of this addendum should be allotted to the particular percentage of correct final digits falling within the correct tens. There are obviously cases where the final digit is correct but the tens digit is incorrect. An examination of the data given in the article of Yerkes and Urban shows that 24.86% of the responses were of the correct tens digit. Therefore 24.86% of the addendum should be added to the percentage of cases in which a correct response occurred; and 24.86% of 5.73% is 1.41%. This percentage, then, should be added to the percentage obtained while the prejudice in favor of 5 or 0 was operative. The percentage of correct responses obtained by Yerkes and Urban was 2.39%; and $2.39\% + 1.41\% = 3.80\%$. This should give us the correct percentage of correct responses obtained when the prejudice referred to is not present. As a matter of fact, in the present experiment, the percentage of correct responses found was 3.88%. The difference of only .08% between the theoretically calculated percentage and the percentage actually obtained is so small that we are justified in considering that the amounts agree in a striking manner.

It may be objected that the present argument assumes that the judgments were distributed purely by chance, whereas the distribution clearly shows a tendency to judge the interval approximately correctly. It is certainly true that the argument rests upon that assumption. But such an assumption is limited to the distribution *within the tens.* In other words, we assume that, while the ability to judge time-intervals is sufficient to enable us to approximate the correct length, this approximation is only to the correct tens digit and not as close as the final digit. Within the tens the distribution of the responses is assumed to be fairly in accordance with the laws of chance. It is possible to discover from our data whether this is actually the case. An

examination of the results gives the following distribution of responses in terms of the final digits within the correct tens.

| Distance of digit from correct digit | —4 | —3 | —2 | —1 | 0 | +1 | +2 | +3 | +4 | +5 |
|---|---|---|---|---|---|---|---|---|---|---|
| Number of cases | 20 | 25 | 30 | 20 | 35 | 41 | 25 | 30 | 38 | 30 |

This shows that the distribution is approximately by chance and is by no means grouped about the correct digit. It is therefore justifiable to assume the chance distribution within the correct tens, which is required by the preceding argument. If the calculations are correct, this close correspondence between the theoretically calculated percentage and the percentage actually observed confirms the accuracy of the present experiment, and shows conclusively that the prejudice mentioned did decidedly affect the judgments of the subjects of Yerkes and Urban's experiment. The difference in the percentage of correct responses as reported by the two investigations is therefore clearly traceable to the difference in method.

Using the mean as a measure of central tendency, Yerkes and Urban found that the 18 sec. interval of "estimation" was overestimated (by 1.5 sec.). In the present investigation the 15 sec. interval in estimation is overestimated also (by 7.27 sec.). In this the two experiments agree. For the longer intervals there is a marked difference. For the 36, 72, and 108 sec. intervals of estimation they found an underestimation of 2.9, 8.8, and 0.2 sec. respectively. I find an *overestimation* in the case of the 30, 60, and 100 sec. intervals of 8.23, 8.55, and 4.31 sec. This difference in the direction of error may be a significant function of the methods used. Table VI gives the means mentioned, as well as the medians and measures of variability. All were calculated from the unclassed measures.

TABLE VI

| Int. in Sec. | Interp. | Mean | $\sigma$ | PEm | Mdn |
|---|---|---|---|---|---|
| 15 | Estimation | 22.27 | 7.79 | .52 | 19.5 |
| 60 | Estimation | 68.55 | 21.40 | 1.44 | 61.1 |
| 100 | Estimation | 104.31 | 26.35 | 1.77 | 100.5 |
| 30 | Estimation | 38.23 | 12.20 | .82 | 35.2 |
| 30 | Prose by Exp. | 37.25 | 12.05 | .81 | 34.8 |
| 30 | Prose by Subj. | 38.36 | 10.89 | .73 | 36.6 |
| 30 | Poetry by Exp. | 39.07 | 12.25 | .82 | 37.3 |
| 30 | Poetry by Subj. | 40.89 | 14.21 | .85 | 38.7 |
| 30 | Dictation | 32.41 | 10.65 | .67 | 30.3 |

In the course of the experiment it was noticed that a subject would occasionally lose all track of the interval, or would be disturbed by some thought or unavoidable distraction, so that certain responses were without doubt abnormally shortened or prolonged. The judgment of 105 sec. for a 30 sec. interval of estimation is certainly not free from some uncontrolled influence. In consideration of this fact, it seems that a measure, such as the median, which does not consider

the value of the extremes, is a more accurate measure of the central tendency than the mean. Using the median as a measure, the 15 and 30 sec. intervals are seen to be overestimated. The 60 and 100 sec. intervals show no definite overestimation if the median is used.

In summing up, the following points are significant:

1. The method of reproduction produces more accurate results than the method of statement in terms of the standard unit. The error of the latter method in the case of Yerkes and Urban's results was theoretically calculated, and the corrected percentage of correct responses is shown to agree with the percentage obtained by the former method.
2. While the percentage of correct responses is greater when the method of reproduction is used, the percentage thus obtained (3.88%) is very low and shows that time-estimation is not very accurate.
3. Results obtained by the method of reproduction do not agree with those obtained by the method of statement in terms of the standard unit in showing that all intervals of greater length than 18 sec. are underestimated. On the contrary, all intervals were overestimated when the method of reproduction was used. If the median is used as the measure of central tendency, the overestimation of the 60 and 100 sec. intervals is not significant. This is true no matter what interpolations are used.
4. Poetry as an interpolation causes the interval to be estimated as longer than an interval with no interpolation (*i. e.*, estimation) or than one with the reading of prose.
5. The interval is estimated as longer if the subject reads the prose or the poetry than if the material is read by the experimenter.
6. Dictation causes the interval to be estimated as shorter than one with any other of the interpolations used, but not necessarily shorter than the actual interval.
7. The results of Yerkes and Urban agree with this study in the comparison of one interpolation with another, as to the effect upon the judgments, although they vary from the present findings in the relation of the judgments of the intervals to the actual intervals.